

Tracking Position and Orientation in a Large Volume

Henry A. Sowizral
Boeing Computer Services
Research and Technology
P.O. Box 24346 MS 7L-22
Seattle, WA 98124-0346
email: henry.sowizral@boeing.com

James C. Barnes
Logitech, Inc.
6607 Kaiser Drive
Fremont, CA 94555

ABSTRACT

Tracking position and orientation accurately and precisely is important in building effective virtual reality systems. Commercially available trackers perform six-degrees-of-freedom tracking adequately but only over limited volumes. Large volume trackers are not available commercially, yet they are necessary to build effective augmented reality¹ systems. This paper presents a new approach to large-volume six-degree-of-freedom tracking. Our approach uses ultrasonic measurement technology and cellular deployment to achieve a relatively low-cost large-volume tracker.

INTRODUCTION

An augmented reality (AR) system [Caudell and Mizell 92] augments a user's view of his surroundings with pictorial, textual, or both kinds of information. AR systems seem very promising for use in simplifying tasks requiring touch labor, especially manufacturing, assembly, and repair tasks. One possible application of AR being studied at Boeing uses pictorial information to help assembly workers build wire bundles more rapidly and accurately. The AR system effectively "draws" the placement of wires on a formboard² by successively drawing each wire's path on a worker's AR display. The worker sees the wire path as if it were resting on the surface of the formboard. The assembly worker removes the appropriate wire from that wire-bundle's assembly kit and attaches that wire to the formboard along the path drawn by the AR system.

Commercially available tracking systems are not adequate for use in most AR applications because workers seldom perform touch labor in a volume that is only five to seven feet across. A long wire bundle may require as many as five adjacent formboards. An AR system capable of supporting such long wire-bundles needs a tracker that operates in a volume at least a three by three by forty feet.

AR systems are not alone in benefiting from large volume tracking. VR applications also benefit. By translating a VR user's movement within a large physical volume into the equivalent movement in the virtual reality, a VR system significantly enhances the illusion of immersion. Gross body movements add qualitatively to the limited kinesthetic feedback produce by moving or rotating the head. While large-volume tracking can enhance the VR experience, it is fundamentally important in AR applications.

Previous Work

The literature documents various methodologies for performing 6DOF tracking including, mechanical [Sutherland 68, Fake Space, Smith 84], magnetic [Ascension, Polhemus], ultrasonic [Logitech], and optical [Antonsson 89, Cook 88, Sorenson 89] approaches. Various surveys [Applewhite 91, Ferrin 91, Meyer *et al.* 92] discuss these tracking technologies in more detail. One approach to large-volume 6DOF tracking is an optical technique successfully demonstrated by the University of North Carolina [Ward 92].

¹ Augmented reality (AR) is also known as see-through virtual reality.

² A formboard is an easle-like structure that hold a wire-bundle during assembly.

The UNC system relies on two-foot-square metal ceiling tiles containing a matrix of light-emitting diodes. The diodes are spaced evenly along a tile's length and width, and are located at a constant depth within the tile. By assembling the tiles into a larger grid, UNC has constructed a large ceiling filled with evenly-spaced light-emitting diodes. The system's user wears a head-mounted display with four lateral-effect "cameras" that "see" a subset of the light-emitting diodes. By appropriately lighting the ceiling and computing which "camera" sees which light emitting diodes, the system can compute the user's position and orientation.

The current UNC ceiling sits 9' above the floor and covers a 10'x12' area. Within this volume, the system can resolve lateral position to within 2mm and rotational orientation to within 0.2 degrees. The system, however, requires a non-trivial superstructure to support the ceiling tiles and the tiles require careful alignment. The four head-mounted, lateral-effect "cameras" are quite heavy and bulky. In general, the UNC tracker does not lend itself to widespread installation since it requires a significant investment to duplicate.

Overview of Our Approach

Ultrasonic technology has a demonstrated cost advantages over other commercially available six-degree-of-freedom (6DOF) tracking technologies. Logitech's ultrasonic system is currently the least expensive commercially available 6DOF tracker.

Our approach uses Logitech's 6DOF ultrasonic tracking technology and cellular deployment to achieve a large-volume 6DOF tracker. Ultrasonic trackers have a constrained operating range, but if we operate multiple trackers in a cellular architecture, we can track over a much larger volume. A cell's ultrasonic transmitter will typically operate only when the detector is within range of that cell's source. By overlapping cells and using simple predictive logic, a computer can switch from cell to cell as the detector moves out of one cell's operating range and into an adjacent cell's operating range.

The Basic Cell

The basic Logitech 6DOF tracker consists of three parts: a source, a detector, and a control unit. The source has three ultrasonic emitters at the vertices of a large plastic triangle. The detector has three microphones at the vertices of a small plastic triangle. The control unit encloses a microprocessor, signal generation hardware, and three cables that respectively connect the control unit to the computer, to the source, and to the detector.

In operation, the control unit starts a timer and sequentially generates an ultrasonic pulse at each emitter. When the pulses reach the detector, the control unit records the time-of-flight information for each source-emitter detector-microphone pair. The control unit then computes the detector's position and orientation by triangulation. In "raw mode," however, the control unit provides the time-of-flight information directly. Raw mode allows users of the ultrasonic tracker to calibrate the source and detector against one another. Such calibration removes variations due to manufacturing and permits the user to write a triangulation algorithm that improves tracker accuracy.

Figure 1 shows a source and a detector. The source's apex emitter lies flat. The base emitters angle upwards 30° from vertical. Each emitter generates a conically shaped ultrasonic field. The unit can track the detector only where the three fields overlap. In addition to the three emitters, the source also contains a microphone mounted just below the triangle's apex and inside the acoustic cones generated by the base emitters. The source-mounted microphone allows the system to compensate for variations in atmospheric conditions. The distance between the base emitters and the apex microphone does not change, but time-of-flight does change with changes in ambient temperature and pressure. The control unit uses the base-emitter to apex-microphone time-of-flight information to translate other time-of-flight measurements into distance.

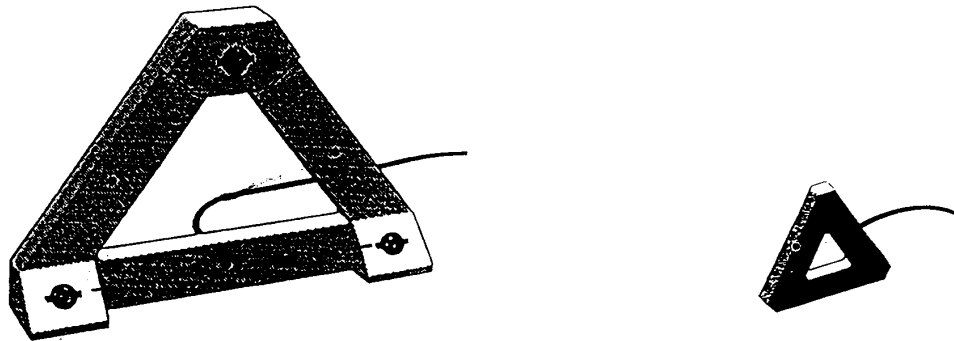


Figure 1. A Logitech source and detector

A cell's geometry and acoustic structure is complex but can be approximated by a truncated cone of constant radius, as shown in Figure 2. The system's tracking region starts approximately six inches below the source and extends four and a half feet further.. The combined acoustic energy of the three emitters spreads out over a 100° field-of-regard. The tracked region's diameter expands by approximately 7 inches for each six inch drop in height. At its base, the cone's tracking radius extends 7 feet 10 inches. By hanging the ultrasonic source seven feet eight inches above the floor, a cell can accommodate individuals 6 foot 6 inch tall and still track a circular region slightly larger than three feet in diameter. In raw mode, a cell's transitional precision is 0.3mm, its accuracy 0.7mm. The rotational precision, 3 feet from the source, is 0.06° , its accuracy 0.15° ; at six feet, rotational precision drops to 0.11° and accuracy drops to 0.3° .

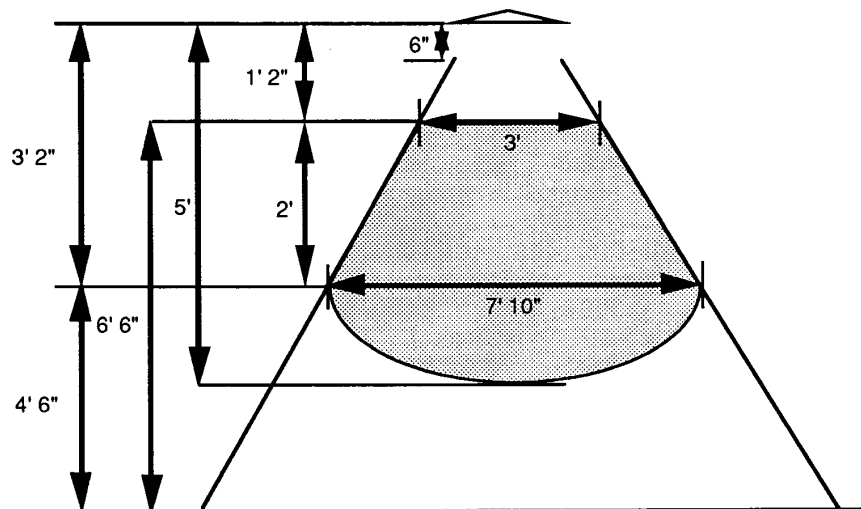


Figure 2. A single cell, its geometry and acoustic structure.

Multiple Cells

Figure 3 shows a profile for a two cell tracker. The geometry is similar to that of the basic cell. Two sources hang nine feet above the ground and three feet apart. Their regions-of-regard touch six feet six inches above the ground and overlap below that level.

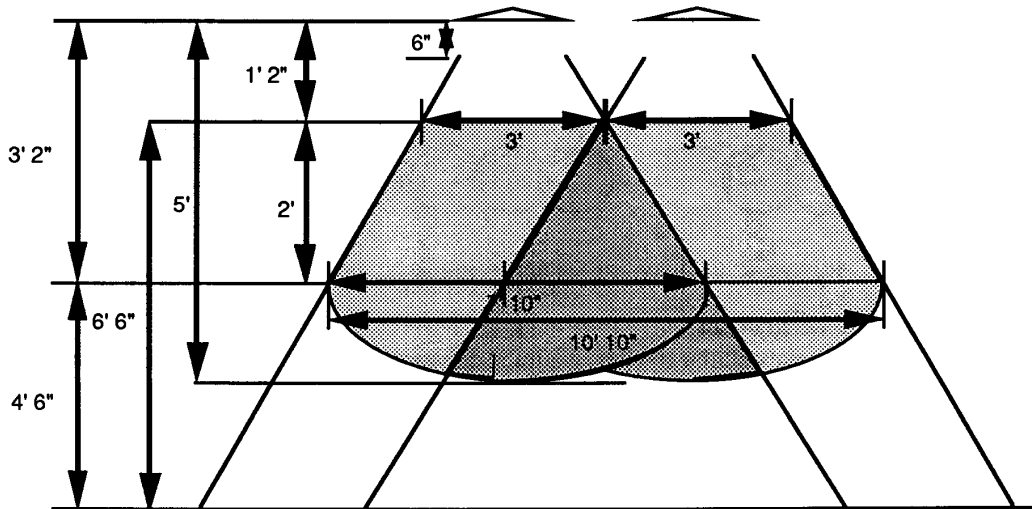


Figure 3. Two cells, their geometry and acoustic structure.

Unfortunately, two adjacent cells cannot operate simultaneously because the cells all emit ultrasonic pulses at the same frequency. Two neighboring cells that operate at the same time will interfere with one another. If we wish to build a cellular tracker, we need to provide a method for selectively activating cells.

We built a switch box that substitutes one source for another under computer control. The switch box plugs into the Logitech controller in place of a source. The sources in turn plug into the switch box. The box contains a set of analog switches ganged together so that they can selectively connect the wires from the Logitech controller to any one source. The prototype switch box can control only two cells, but the technique easily accommodates more.

Volume filling

Figure 4 shows a plan view of various basic cell arrangement that result in large area coverage. If r represents a single cell's detection radius then we can cover the greatest volume by separating the sources by a distance of $\frac{r\sqrt{3}}{2}$. In our example, r is 1' 6", if we separate the cell centers by 1' 4" we achieve maximal coverage. Two cells arranged in this manner cover a maximal lineal distance of 5' 8". Four cells can cover a rhombic region 5' 8" by 5' 8". An arrangement of 20 cells (five rows of four cells) permits the coverage of an area 11' 2" x 14', exceeding the tracking region of UNC's tracker.

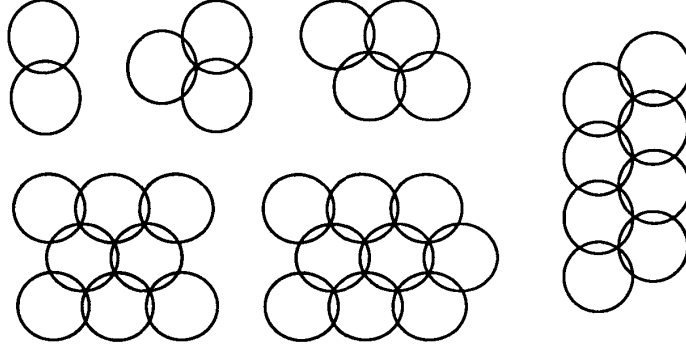


Figure 4. Cell layouts for achieving large area coverage.

Alignment and Calibration

Before using the cellular tracker, we must construct a single coordinate frame across all cells. One technique is to assemble the sources so they all lie carefully oriented and in a single plane. Then, by physically measuring the translational offsets between sources, we can construct a common coordinate frame. This process of assembly and measurement is, of necessity, an exacting and labor intensive process prone to introducing fairly large systematic errors.

Another approach is to use the partially working tracker itself to measure the pairwise relative position and orientation of adjacent cells. From the pairwise positions and orientations, we can construct a common coordinate frame. We use this latter approach.

We first need to define some basic terms and symbols before discussing how we construct the common coordinate frame. We use D_C to represent the position and orientation of the detector relative to a particular cell's coordinate frame. The matrix D_C encodes orientation as three direction cosines and position as a three vector displacement. We use the notation T_{AB} to represent a transformation from coordinate system B into coordinate system A. By evaluating:

$$D_R = T_{RC} \cdot D_C$$

we translate the position-orientation matrix D_C from cell C's coordinate frame into cell R's coordinate frame.

Static Alignment

We begin static alignment by first assigning a unique number to each cell. We choose one cell, say C_1 , as the cell that defines the cellular tracker's coordinate frame. We place our detector in the overlap region between the root cell (C_1) and an adjacent cell (C_2) and measure the detector's position and orientation with respect to the root cell (D_{C_1}) and the adjacent cell (D_{C_2}). From these two measurements we compute a transform matrix that can transform measurements, made relative to one cell, into measurements relative to the other cell. That transformation matrix is computed as follows:

$$T_{C_1 C_2} = D_{C_1} \cdot D_{C_2}^{-1}$$

We associate this matrix with cell C_2 . Thereafter, before that cell returns a position-orientation measurement, it pre multiplies the measured position-orientation matrix with its stored transformation matrix.

We repeat this alignment process for each of the remaining cells. The more general problem of registering the tracker's coordinate system with the virtual world's coordinate system and the physical world's coordinate system is beyond the scope of this paper see [Janin, Mizell and Caudell 93] for a description of the process.

Cell Switching

Deciding when to switch from one cell to another requires carefully tracking the detector's position over time. From time-indexed position information, we can compute the detector's velocity. Similarly, we can compute the detector's acceleration from time-indexed velocities. Knowing the detector's position, velocity, and acceleration we can predict when the detector will move from being mainly inside one cell to being mainly inside an adjacent cell. The cell switching decision must include a slight amount of hysteresis to prevent oscillations.

System Precision and Accuracy

The system's precision and accuracy reduce as the square root of the distance (number of cells) from the root cell. Thus, assuming our rectangular arrangement of 20 cells (five rows of four cells) that cover a region 11' 2" x 14' and assuming we choose one corner as the system's root coordinate system, the translational precision along the column drops to $0.3\sqrt{4}$ mm or 0.6mm, the translation precision along the row drops to $0.3\sqrt{5}$ or 0.67mm, and the translation precision along the diagonal drops to $0.3\sqrt{6}$ or 0.73mm. Worst case (the diagonal) the tracker's translation precision is 0.73mm with an accuracy of 1.7mm; its rotational precision at 3 feet is 0.14° with an accuracy of 0.36° and at 6 feet its rotation precision drops to 0.28° with an accuracy of 0.72° .

By choosing the central cell as the system's root cell, we can improve worst-case precision and accuracy because we reduce the cell-distance from the root cell. The furthest cell is a cell distance of 4 cells from the root resulting in a system error a factor of $\sqrt{4}$ worse than a single cell. Thus, translation precision is 0.6mm with an accuracy of 0.14mm and rotational precision at 3 feet is 0.12° with an accuracy of 0.3° and at 6 feet rotational precision is 0.22° with an accuracy of 0.6° .

Dynamic alignment

We can alternatively perform on-the-fly or dynamic cell alignment. Dynamic alignment is adapted directly from the static alignment technique. Whenever a detector is in range of two or more cells, we can measure the detector's position-orientation with respect to one cell and then, in the next time interval, with respect to the adjacent cell. These two measurements permit us to construct a transform matrix.

Though the dynamic alignment technique is conceptually quite simple, its implementation requires considerable care. We cannot make two simultaneous measurements of the detector from two adjacent cells and rarely will a detector stay at the same position-orientation over the interval of time needed to take the two consecutive measurements. This requires the system to track the detector's trend information fairly carefully, specifically position-orientations, velocities, and accelerations. Using this information and knowing the measurement rate, we can predict a detector's future position-orientation at the time that we take the second measurement. We can improve the transformation matrix iteratively by continually switching between cells whenever the detector is in a cell overlap region and recomputing the transform. Admittedly, a dynamically aligned cellular tracker will not have the precision or accuracy of a statically aligned tracker.

Multiple Detectors

The Logitech controller can track up to four detectors simultaneously with one source. The cellular system can also track multiple detectors, if all the detectors are in the same cell. As an example, a cell can easily track a user's head and right arm at the same time. However, if the user's head stays inside one cell while his arm moves into an

adjoining cell, we have a problem. Either we need to stop tracking the user's arm or we need to use a method for tracking detectors in adjoining cells. In applications where we can feasibly stop tracking one detector our problem reduces to deciding which detector takes precedence. In applications where we must continue tracking both detectors, we can do so by time-division multiplexing adjacent cells. In the two cell case, we track one detector inside the first cell at even time intervals and we track the second detector inside the second cell at odd time intervals. If necessary, time-division multiplexing can work across more cells; however, not without a price. Time-division multiplexing reduces a cell's measurement rate. If the tracker is multiplexing two cells, update rates are halved, dropping from 50Hz to 25Hz. Multiplexing across three cells reduces the measurement rate to 16Hz and four cells reduces the rate to 12Hz—a barely acceptable measurement rate.

Conclusions and Future Directions

An AR system requires high-precision tracking so that it can position its virtual objects accurately within the physical surroundings and make those objects appear to remain stationary even when the user moves or rotates his head. A measurement error of a single degree can displace a virtual object by a substantial amount. Positioning objects accurately with respect to the physical surroundings requires good registration between the physical coordinate frame and the 6DOF tracker's coordinate frame. Keeping virtual objects from spurious displacements requires the 6DOF tracker to perform its measurements precisely.

We have presented a design for a low-cost, large-volume, reasonably high-precision and high-accuracy cellular tracker. The stringent requirements of AR applications provided the main motivation for the tracking system. Though the system provides appropriate precision and accuracy for use in laboratory and less demanding prototype applications, we must continue to develop trackers that have higher precision and higher accuracies and work in large volumes.

We may be able to extract better resolution and accuracy by enlarging the separation between a source's three ultrasonic emitters. By enlarging the separating among the emitters we enlarge the system's baseline allowing a more accurate triangulation computation. It appear possible to enlarging the source's baseline without physically modifying the source triangles. Instead, we can modify the switch box so that two emitters from one source operate in conjunction with an emitter from an adjacent cell's source.

Acknowledgments

We would like to thank David Mizell and Thomas Caudell who began the AR effort at Boeing and without whom this problem would not have been posed; Michael Deering of Sun Microsystems, Inc. for catalyzing this work; Thomas King for measuring a Logitech tracker's precision and accuracy when operated in raw mode; Victor Fong, Ali Moayer and Laurent Piguet for working on the design and test of the switch box; and, Chris Esposito, Adam Janin, and Karel Zikan for their discussions and insight into large volume tracking.

References

- Antonsson, E.K., and R.W. Mann. "Automatic 6-D.O.F. Kinematic Trajectory Acquisition and Analysis. *Journal Dynamic Systems, Measurement, and Control*, 111, (March 1989) pp.31-39.
- Applewhite H., "Position Tracking in Virtual Reality," in *Proceedings of the Second Annual Virtual Reality Conference and Exhibition*, S.K. Helsel(ed.), September 23-25, 1991. Meckler Corporation, Westport, CT
- Caudell, Thomas P., and David W. Mizell, "Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Process," in proceeding Hawaii International Conference on Systems Science, 1991, pp. 659-669.

Cook, Anthony. "The Helmet -Mounted Visual System in Flight Simulation: Recent developments in technology and use," in proceedings *Royal Aeronautical Society*, London, England, April 12-13, 1988, pp. 214-232.

Fake Space Labs Incorporated. Binocular Omni-Orientation Monitor (BOOM), Menlo Park, CA.

Ferrin, Frank J., "Survey of Helmet Tracking Technologies," in proceeding *SPIE Large-Screen-Projection, Avionic, and Helmet-Mounted Displays* (1991), Vol. 1456, pp. 86-94.

Janin, Adam L., David W. Mizell, and Thomas P. Caudell, "Calibration of Head-Mounted Displays for Augmented Reality Applications," submitted VRAIS 93.

Meyer, Kenneth, Hugh L. Applewhite, and Frank A. Biocca, "A Survey of Position Trackers," *Presence*, Vol 1, Number 2, Spring 1992, pp. 173-200.

Smith Jr. B. R. "Digital Head Tracking and Position Prediction for Helmet Mounted Visual Display Systems," in proceedings *AIAA 22nd Aerospace Meeting*, Reno, NV, Jan 9-12, 1984.

Sorensen, Brett, Max Donath, Guo-Ben Yang, and Roland Starr. "The Minnesota Scanner: a Prototype Sensor for Three-dimensional Tracking of Moving Body Segments," *IEEE Transactions on Robotics and Automation*, vol 5., issue 4, (August 1989), pp. 499-509.

Sutherland, Ivan. "A Head-Mounted Three Dimensional Display," in proceedings *1968 Fall Joint Computer Conference*, AFIPS Conference Proceedings, 33 (1968) pp. 757-764.

Ward, Mark, Ronald Azuma, Robert Bennett, Stefan Gottschalk, and Henry Fuchs, "A Demonstrated Optical Tracker With Scalable Work Area for Head-Mounted Display Systems," in proceedings *1992 Symposium on Interactive 3D Graphics* (29 March-1 April 1992), Cambridge, MA, pp. 43-52.